What kind of tensors are compressible?

Tianyi Shi

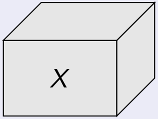
Cornell University

ts777@cornell.edu

June 28, 2019

Work with: Alex Townsend (Cornell University)

What is a tensor?

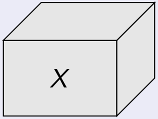


Why are low rank tensors important?

Explicit storage (3D):

$$\prod_{i=1}^{3} n_i.$$

What is a tensor?



Why are low rank tensors important?

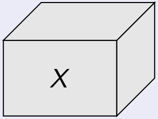
Explicit storage (3D):

$$\prod_{i=1}^{3} n_i.$$

Methodologies to understand the compressibility of tensors:

- Algebraic structures: $\mathcal{X}_{i,j,k} = f(x_i, y_j, z_k)$

What is a tensor?



Why are low rank tensors important?

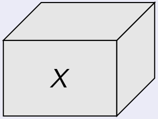
Explicit storage (3D):

$$\prod_{i=1}^{3} n_i.$$

Methodologies to understand the compressibility of tensors:

- Algebraic structures: $\mathcal{X}_{i,j,k} = f(x_i, y_j, z_k)$
- Smoothness: $f(x, y, z) \approx p_n(x, y, z)$

What is a tensor?



Why are low rank tensors important?

Explicit storage (3D):

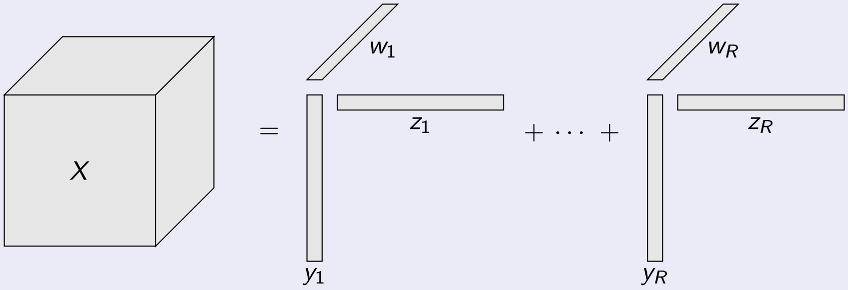
$$\prod_{i=1}^{3} n_i.$$

Methodologies to understand the compressibility of tensors:

- Algebraic structures: $\mathcal{X}_{i,j,k} = f(x_i, y_j, z_k)$
- Smoothness: $f(x,y,z) \approx p_n(x,y,z)$
- **Displacement structure**

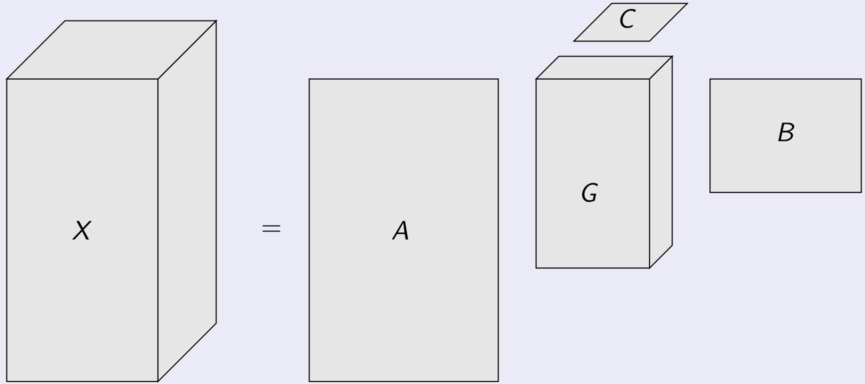
Tensor decompositions

CP decomposition [Kolda & Bader, 09]



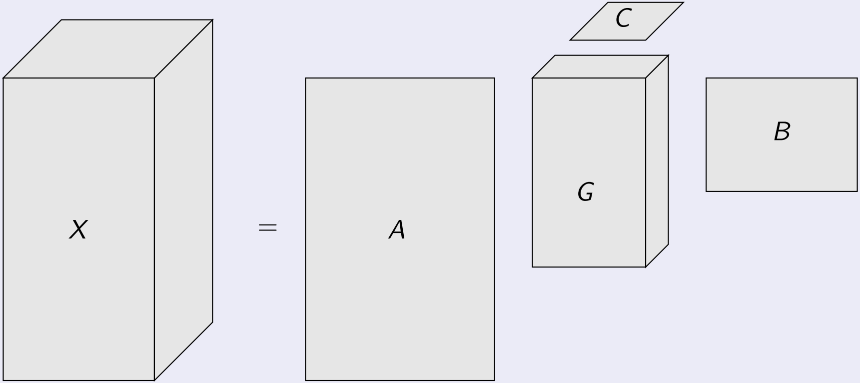
Tensor decompositions

Tucker decomposition [Tucker, 1963]



Tensor decompositions

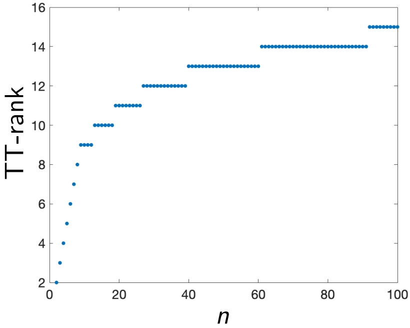
Tucker decomposition [Tucker, 1963]

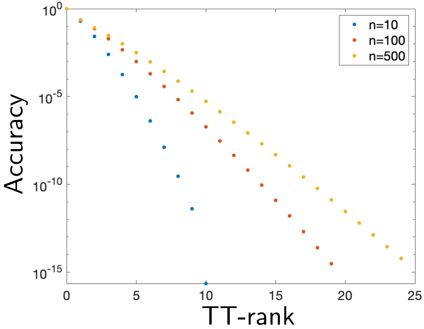


Tensor-train (TT) decomposition

Example: Hilbert tensor

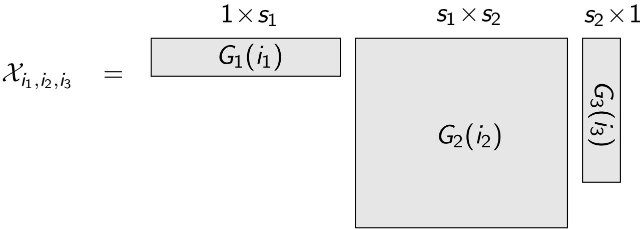
$$\mathcal{H}_{i,j,k} = \frac{1}{i+j+k-2}, \qquad 1 \le i,j,k \le n.$$





E.g. $n = 100, \epsilon = 10^{-10}$, instead of 100^3, in tensor-train: 25500.

Tensor-train decomposition [Oseledets, 11]

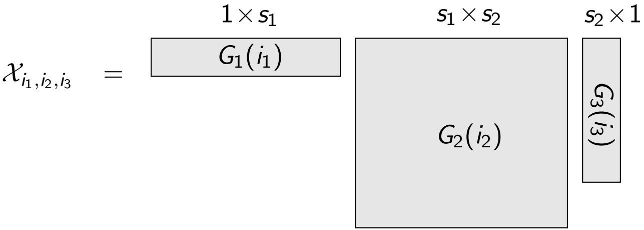


$$\mathrm{rank}^{TT}(\mathcal{X}) = (1, s_1, s_2, 1).$$

Storage:

$$\sum_{k=1}^{3} s_{k-1} s_k n_k .$$

Tensor-train decomposition [Oseledets, 11]



$$\mathrm{rank}^{TT}(\mathcal{X}) = (1, s_1, s_2, 1).$$

Storage:

$$\sum_{k=1}^{3} s_{k-1} s_k n_k .$$

Bound:

$$s_k \leq \mathrm{rank}(X_k), \qquad X_k = \mathrm{reshape}(\mathcal{X}, \prod_{s=1}^{k} n_s, \prod_{s=k+1}^{3} n_s).$$

Numerical tensor-train rank

Numerical tensor-train rank

$\mathrm{rank}_{\epsilon}^{\mathrm{TT}}(\mathcal{X}) = \boldsymbol{s}_{\epsilon}$ where $\boldsymbol{s}_{\epsilon}$ is the smallest vector such that $\mathrm{rank}^{\mathrm{TT}}(\tilde{\mathcal{X}}) = \boldsymbol{s}_{\epsilon}$, $\|\mathcal{X} - \tilde{\mathcal{X}}\|_F \leq \epsilon \|\mathcal{X}\|_F$, and $||\mathcal{X}||_F = \left(\sum_{i_1,i_2,i_3} (\mathcal{X}_{i_1,i_2,i_3})^2\right)^{1/2}$.

Lexicographical ordering

A vector $\boldsymbol{x} = (x_1, \ldots, x_d)$ is less than $\boldsymbol{y} = (y_1, \ldots, y_d)$, denoted by $\boldsymbol{x} <_{\mathrm{lex}} \boldsymbol{y}$, if in the first entry for which the vectors differ, $x_j < y_j$.
In addition, $\boldsymbol{x} \leq_{\mathrm{lex}} \boldsymbol{y}$ if $\boldsymbol{x} <_{\mathrm{lex}} \boldsymbol{y}$ or $x_j = y_j$ for all j.

Displacement structure

Matrix

$$AX + XB^T = G, \qquad A \in \mathbb{C}^{m\times m}, \quad B \in \mathbb{C}^{n\times n},$$

Displacement structure

Matrix

$$AX + XB^T = G, \qquad A \in \mathbb{C}^{m\times m}, \quad B \in \mathbb{C}^{n\times n},$$

3D Tensor

$$\mathcal{X} \times_1 A^{(1)} + \mathcal{X} \times_2 A^{(2)} + \mathcal{X} \times_3 A^{(3)} = \mathcal{G}, \qquad A^{(k)} \in \mathbb{C}^{n_k\times n_k},$$

The k-mode product

For a tensor $\mathcal{X} \in \mathbb{C}^{n_1\times\cdots\times n_d}$ and a matrix $A \in \mathbb{C}^{n_k\times n_k}$

$$(\mathcal{X} \times_k A)_{i_1,\ldots,i_{k-1},j,i_{k+1},\ldots,i_d} = \sum_{i_k=1}^{n_k} \mathcal{X}_{i_1,\ldots,i_d} A_{j,i_k}.$$

Displacement structure

Matrix

$$AX + XB^T = G, \qquad A \in \mathbb{C}^{m\times m}, \quad B \in \mathbb{C}^{n\times n},$$

3D Tensor

$$\mathcal{X} \times_1 A^{(1)} + \mathcal{X} \times_2 A^{(2)} + \mathcal{X} \times_3 A^{(3)} = \mathcal{G}, \qquad A^{(k)} \in \mathbb{C}^{n_k\times n_k},$$

The k-mode product

For a tensor $\mathcal{X} \in \mathbb{C}^{n_1\times\cdots\times n_d}$ and a matrix $A \in \mathbb{C}^{n_k\times n_k}$

$$(\mathcal{X} \times_k A)_{i_1,\ldots,i_{k-1},j,i_{k+1},\ldots,i_d} = \sum_{i_k=1}^{n_k} \mathcal{X}_{i_1,\ldots,i_d} A_{j,i_k}.$$

Matrix (again)

$$X \times_1 A + X \times_2 B = G$$

Rank bound of matrices with displacement structure

If

- A and B are normal matrices,
- $\Lambda(A) \subseteq E$ and $\Lambda(B) \subseteq F$,

then $AX - XB^T = G$, $\mathrm{rank}(G) = \nu$ implies

2-norm [Beckermann & Townsend, 19]

$$\|X - X_{\nu k}\|_2 \leq Z_k(E, F)\|X\|_2.$$

Frobenius norm [S. & Townsend]

$$\|X - X_{\nu k}\|_F \leq Z_k(E, F)\|X\|_F.$$

Zolotarev number [Zolotarev, 1877]

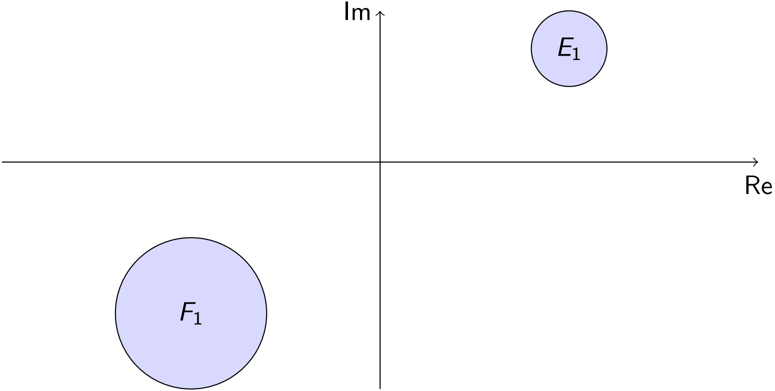
$$Z_k(E,F) := \inf_{r\in\mathcal{R}_{k,k}} \frac{\sup_{z\in E}|r(z)|}{\inf_{z\in F}|r(z)|}, \qquad k \geq 0,$$

E and F are disjoint complex sets and $\mathcal{R}_{k,k}$ is the set of irreducible rational functions of the form $p(x)/q(x)$ with polynomials p and q of degree at most k.

Zolotarev number [Zolotarev, 1877]

$$Z_k(E,F) := \inf_{r\in\mathcal{R}_{k,k}} \frac{\sup_{z\in E}|r(z)|}{\inf_{z\in F}|r(z)|}, \qquad k \geq 0,$$

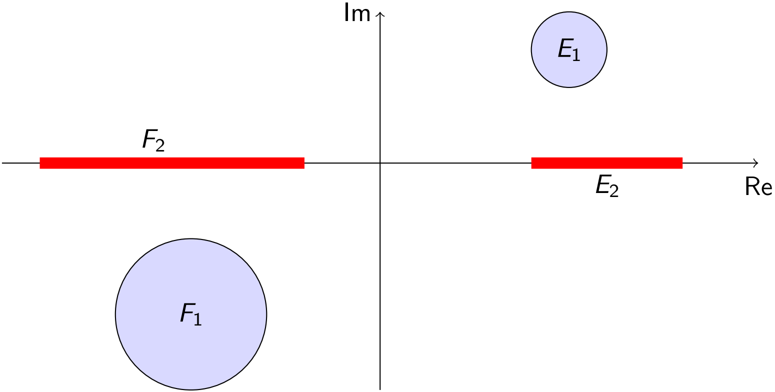
E and F are disjoint complex sets and $\mathcal{R}_{k,k}$ is the set of irreducible rational functions of the form $p(x)/q(x)$ with polynomials p and q of degree at most k.



Zolotarev number [Zolotarev, 1877]

$$Z_k(E,F) := \inf_{r\in\mathcal{R}_{k,k}} \frac{\sup_{z\in E}|r(z)|}{\inf_{z\in F}|r(z)|}, \qquad k \geq 0,$$

E and F are disjoint complex sets and $\mathcal{R}_{k,k}$ is the set of irreducible rational functions of the form $p(x)/q(x)$ with polynomials p and q of degree at most k.



Rank bound of tensors with displacement structure

Minkowski sum separated

For normal matrices $A^{(1)}, A^{(2)}, A^{(3)}$, and disjoint sets E_j and F_j,

$$\Lambda(A^{(1)}) \subseteq E_1, \quad -(\Lambda(A^{(2)}) + \Lambda(A^{(3)})) \subseteq F_1,$$

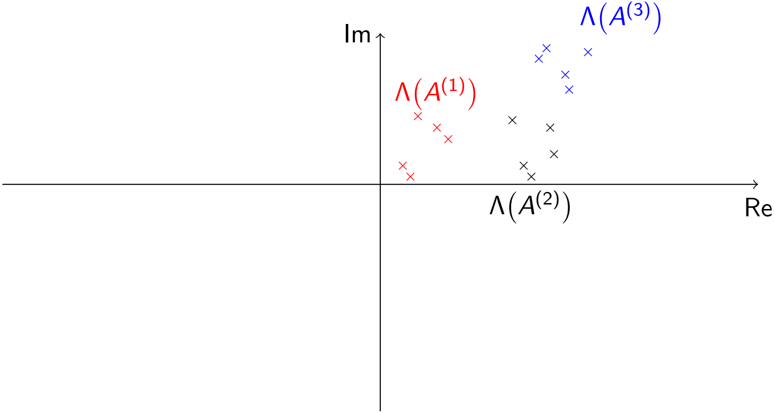
$$\Lambda(A^{(1)}) + \Lambda(A^{(2)}) \subseteq E_2, \quad -\Lambda(A^{(3)}) \subseteq F_2.$$

Minkowski sum separated

$$\Lambda(A^{(1)}) \subseteq E_1, \quad -(\Lambda(A^{(2)}) + \Lambda(A^{(3)})) \subseteq F_1, \quad E_1 \cap F_1 = \varnothing,$$
$$\Lambda(A^{(1)}) + \Lambda(A^{(2)}) \subseteq E_2, \quad -\Lambda(A^{(3)}) \subseteq F_2, \quad E_2 \cap F_2 = \varnothing.$$

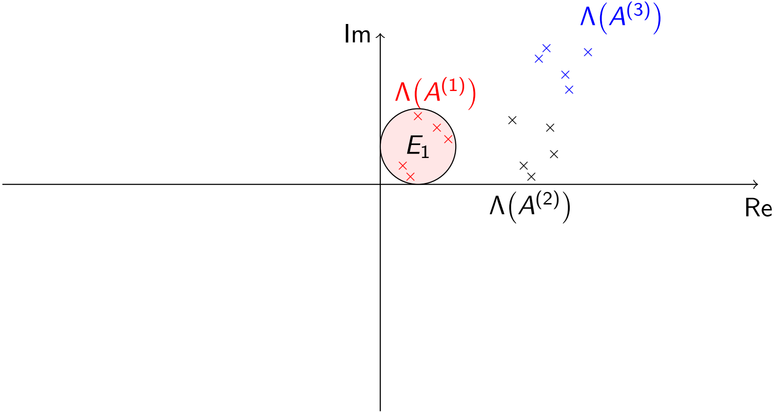
Minkowski sum separated

$$\Lambda(A^{(1)}) \subseteq E_1, \quad -(\Lambda(A^{(2)}) + \Lambda(A^{(3)})) \subseteq F_1, \quad E_1 \cap F_1 = \varnothing,$$
$$\Lambda(A^{(1)}) + \Lambda(A^{(2)}) \subseteq E_2, \quad -\Lambda(A^{(3)}) \subseteq F_2, \quad E_2 \cap F_2 = \varnothing.$$



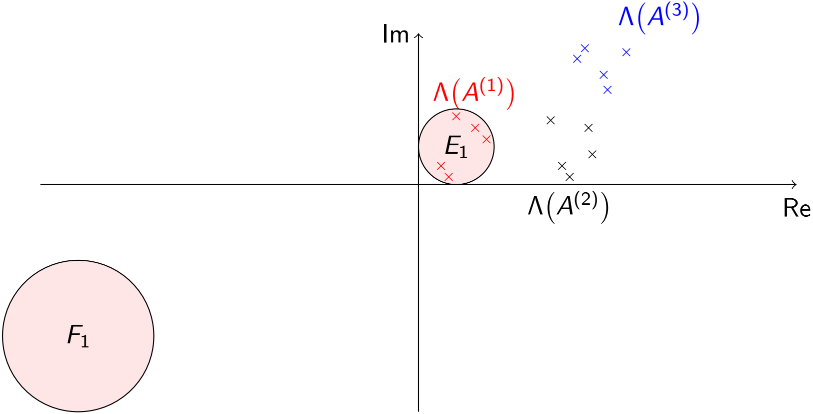
Minkowski sum separated

$$\Lambda(A^{(1)}) \subseteq E_1, \quad -(\Lambda(A^{(2)}) + \Lambda(A^{(3)})) \subseteq F_1, \quad E_1 \cap F_1 = \varnothing,$$
$$\Lambda(A^{(1)}) + \Lambda(A^{(2)}) \subseteq E_2, \quad -\Lambda(A^{(3)}) \subseteq F_2, \quad E_2 \cap F_2 = \varnothing.$$



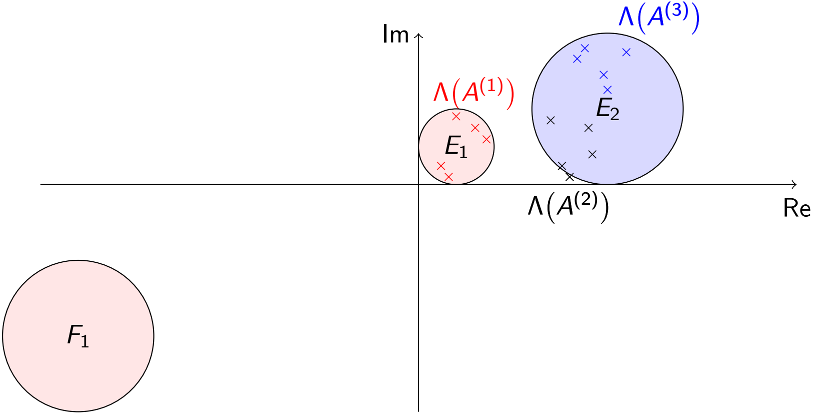
Minkowski sum separated

$$\Lambda(A^{(1)}) \subseteq E_1, \quad -(\Lambda(A^{(2)}) + \Lambda(A^{(3)})) \subseteq F_1, \quad E_1 \cap F_1 = \varnothing,$$
$$\Lambda(A^{(1)}) + \Lambda(A^{(2)}) \subseteq E_2, \quad -\Lambda(A^{(3)}) \subseteq F_2, \quad E_2 \cap F_2 = \varnothing.$$



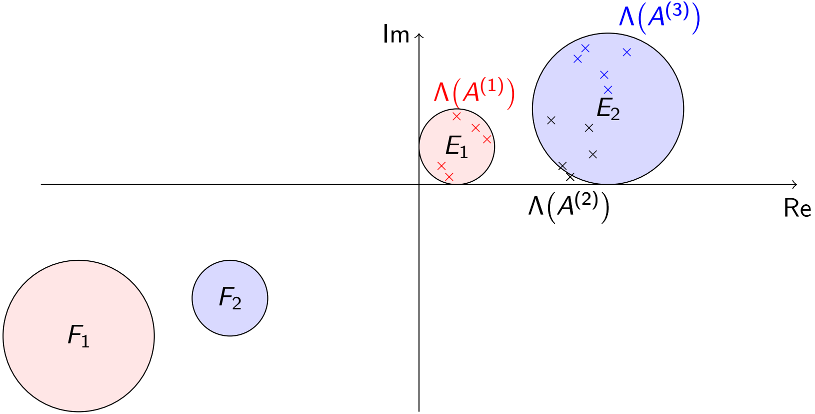
Minkowski sum separated

$$\Lambda(A^{(1)}) \subseteq E_1, \quad -(\Lambda(A^{(2)}) + \Lambda(A^{(3)})) \subseteq F_1, \quad E_1 \cap F_1 = \varnothing,$$
$$\Lambda(A^{(1)}) + \Lambda(A^{(2)}) \subseteq E_2, \quad -\Lambda(A^{(3)}) \subseteq F_2, \quad E_2 \cap F_2 = \varnothing.$$



Minkowski sum separated

$$\Lambda(A^{(1)}) \subseteq E_1, \quad -(\Lambda(A^{(2)}) + \Lambda(A^{(3)})) \subseteq F_1, \quad E_1 \cap F_1 = \varnothing,$$
$$\Lambda(A^{(1)}) + \Lambda(A^{(2)}) \subseteq E_2, \quad -\Lambda(A^{(3)}) \subseteq F_2, \quad E_2 \cap F_2 = \varnothing.$$



Rank bound of tensors with displacement structure (ctd.)

Theorem (S. & Townsend)

Suppose $\mathcal{X} \times_1 A^{(1)} + \mathcal{X} \times_2 A^{(2)} + \mathcal{X} \times_3 A^{(3)} = \mathcal{G}$, *where* $A^{(1)}, A^{(2)}, A^{(3)}$ *are Minkowski sum separated with disjoint sets* E_j *and* F_j *for* $j = 1, 2$. *Then, for a fixed* $0 < \epsilon < 1$, *we have*

$$\mathrm{rank}_{\epsilon}^{\mathrm{TT}}(\mathcal{X}) \leq_{\mathrm{lex}} (1, k_1\nu_1, k_2\nu_2, 1), \quad \nu_j = \mathrm{rank}(G_j), \quad j = 1, 2,$$

where G_j *is the* j*th unfolding of* $\mathcal{G}$ *and* k_j *is an integer so that* $Z_{k_j}(E_j, F_j) \leq \epsilon/\sqrt{3}$.

Rank bound of tensors with displacement structure (ctd.)

Theorem (S. & Townsend)

Suppose $\mathcal{X} \times_1 A^{(1)} + \mathcal{X} \times_2 A^{(2)} + \mathcal{X} \times_3 A^{(3)} = \mathcal{G}$, where $A^{(1)}, A^{(2)}, A^{(3)}$ are Minkowski sum separated with disjoint sets E_j and F_j for $j = 1, 2$. Then, for a fixed $0 < \epsilon < 1$, we have

$$\mathrm{rank}_{\epsilon}^{\mathrm{TT}}(\mathcal{X}) \leq_{\mathrm{lex}} (1, k_1\nu_1, k_2\nu_2, 1), \quad \nu_j = \mathrm{rank}(G_j), \quad j = 1, 2,$$

where G_j is the jth unfolding of $\mathcal{G}$ and k_j is an integer so that $Z_{k_j}(E_j, F_j) \leq \epsilon/\sqrt{3}$.

Special case

If $\Lambda(A^{(j)}) \subseteq [a, b]$ for $0 < a < b < \infty$, and $\gamma_j = \frac{(3a+j(b-a))(3b-j(b-a))}{9ab}$, then

$$\mathrm{rank}_{\epsilon}^{\mathrm{TT}}(\mathcal{X}) \leq_{\mathrm{lex}} (1, k_1\nu_1, k_2\nu_2, 1), \qquad k_j = \left\lceil \frac{\log(16\gamma_j)\log(4\sqrt{3}/\epsilon)}{\pi^2} \right\rceil.$$

Hilbert tensor revisited

$$\mathcal{H}_{i,j,k} = \frac{1}{i+j+k-2}, \qquad 1 \le i, j, k \le n.$$

Hilbert tensor revisited

$$\mathcal{H}_{i,j,k} = \frac{1}{i+j+k-2}, \qquad 1 \leq i,j,k \leq n.$$

$$\mathcal{H} \times_1 D + \mathcal{H} \times_2 D + \mathcal{H} \times_3 D = \mathcal{S},$$

$\mathcal{S}$ is the tensor of all ones and D is a diagonal matrix with $D_{ii} = i - \frac{2}{3}$.

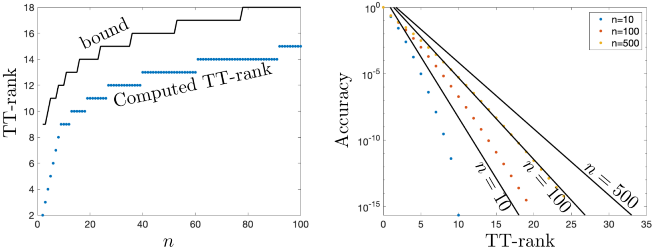
Hilbert tensor revisited

$$\mathcal{H}_{i,j,k} = \frac{1}{i+j+k-2}, \qquad 1 \le i,j,k \le n.$$

$$\mathcal{H} \times_1 D + \mathcal{H} \times_2 D + \mathcal{H} \times_3 D = \mathcal{S},$$

$\mathcal{S}$ is the tensor of all ones and D is a diagonal matrix with $D_{ii} = i - \frac{2}{3}$.

$$\mathrm{rank}_\epsilon^{\mathrm{TT}}(\mathcal{H}) \le_{\mathrm{lex}} (1, s_1, s_1, 1), \qquad s_1 = \left\lceil \frac{1}{\pi^2} \log\left(\frac{16n(2n-1)}{3n-2}\right) \log\left(\frac{4\sqrt{3}}{\epsilon}\right) \right\rceil.$$



Example: Solution to 3D Poisson equation

$$-(u_{xx} + u_{yy} + u_{zz}) = f \text{ on } \Omega = [-1, 1]^3, \qquad u|_{\partial\Omega} = 0.$$

$$\mathcal{X}_{i,j,k} = u(x_i, y_j, z_k), \quad \mathcal{F}_{i,j,k} = f(x_i, y_j, z_k),$$

Example: Solution to 3D Poisson equation

$$-(u_{xx} + u_{yy} + u_{zz}) = f \text{ on } \Omega = [-1, 1]^3, \qquad u|_{\partial\Omega} = 0.$$

$$\mathcal{X}_{i,j,k} = u(x_i, y_j, z_k), \quad \mathcal{F}_{i,j,k} = f(x_i, y_j, z_k),$$

$$\mathcal{X} \times_1 K + \mathcal{X} \times_2 K + \mathcal{X} \times_3 K = \mathcal{F}, \qquad K = -\frac{1}{h^2} \begin{bmatrix} 2 & -1 & & \\ -1 & \ddots & \ddots & \\ & \ddots & \ddots & -1 \\ & & -1 & 2 \end{bmatrix},$$

Example: Solution to 3D Poisson equation

$$-(u_{xx} + u_{yy} + u_{zz}) = f \text{ on } \Omega = [-1, 1]^3, \qquad u|_{\partial\Omega} = 0.$$

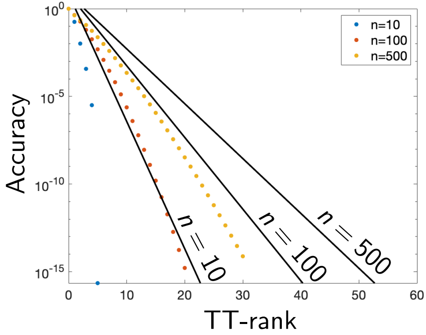
$$\mathcal{X}_{i,j,k} = u(x_i, y_j, z_k), \quad \mathcal{F}_{i,j,k} = f(x_i, y_j, z_k),$$

$$\mathcal{X} \times_1 K + \mathcal{X} \times_2 K + \mathcal{X} \times_3 K = \mathcal{F}, \qquad K = -\frac{1}{h^2} \begin{bmatrix} 2 & -1 & & \\ -1 & \ddots & \ddots & \\ & \ddots & \ddots & -1 \\ & & -1 & 2 \end{bmatrix},$$

Consider $f = 1$,

$$\mathrm{rank}_\epsilon^{\mathrm{TT}}(\mathcal{X}) \leq_{\mathrm{lex}} (1, s_1, s_1, 1), \quad s_1 = \left\lceil \frac{1}{\pi^2} \log\left(\frac{16(n^2+2)(2n^2+1)}{9n^2}\right) \log\left(\frac{4\sqrt{3}}{\epsilon}\right) \right\rceil.$$

Example: Solution of 3D Poisson equation



3D Poisson solver

- Constructive bound proof
- Solve with ultraspherical spectral methods [Fortunato & Townsend, 17]
- Solve in TT format super fast with complexity $\mathcal{O}(n(\log n)^2(\log 1/\epsilon)^2)$

Takeaways

- Several methodologies guarantee compressibility of tensors.
- Numerical TT ranks of tensors with specific displacement structure is $\mathcal{O}(\log n \log(1/\epsilon))$.
- 3D Poisson solver with optimal complexity and spectral accuracy in tensor-train format.

Ongoing work

- Make the fast Poisson solver open-source codes.
- Ranks of matrices/tensors with more general displacement structure.
- Fast solve other separable PDEs (e.g. generalized Sylvester equation).